

How to Train Your Energy-Based Models

Year: 2021 | Citations: 295 | Authors: Yang Song, Diederik P. Kingma

Abstract

Energy-Based Models (EBMs), also known as non-normalized probabilistic models, specify probability density or mass functions up to an unknown normalizing constant. Unlike most other probabilistic models, EBMs do not place a restriction on the tractability of the normalizing constant, thus are more flexible to parameterize and can model a more expressive family of probability distributions. However, the unknown normalizing constant of EBMs makes training particularly difficult. Our goal is to provide a friendly introduction to modern approaches for EBM training. We start by explaining maximum likelihood training with Markov chain Monte Carlo (MCMC), and proceed to elaborate on MCMC-free approaches, including Score Matching (SM) and Noise Contrastive Estimation (NCE). We highlight theoretical connections among these three approaches, and end with a brief survey on alternative training methods, which are still under active research. Our tutorial is targeted at an audience with basic understanding of generative models who want to apply EBMs or start a research project in this direction.